JB Journals.ASM.org

In Vitro Evolution of an Archetypal Enteropathogenic *Escherichia coli* Strain

Shahista Nisa,[a] Tracy H. Hazen,[b] Lillian Assatourian,[a] Jean-Philippe Nougayrède,[a]* David A. Rasko,[b] Michael S. Donnenberg[a]

Department of Medicine[a] and Institute for Genome Sciences and Department of Microbiology and Immunology,[b] School of Medicine, University of Maryland, Baltimore, Maryland, USA

Enteropathogenic *Escherichia coli* (EPEC) is a leading cause of infantile diarrhea in developing countries. EPEC strain E2348/69 is used worldwide as a prototype to study EPEC genetics and disease. However, isolates of E2348/69 differ phenotypically, reflecting a history of *in vitro* selection. To identify the genomic and phenotypic changes in the prototype strain, we sequenced the genome of the nalidixic acid-resistant (Nal^r) E2348/69 clone. We also sequenced a recent *nleF* mutant derived by one-step PCR mutagenesis from the Nal^r strain. The sequencing results revealed no unintended changes between the mutant and the parent strain. However, loss of the pE2348-2 plasmid and 3 nonsynonymous mutations were found in comparison to the published streptomycin-resistant (Str^r) E2348/69 reference genome. One mutation is a conservative amino acid substitution in *ftsK*. Another, in *gyrA*, is a mutation known to result in resistance to nalidixic acid. The third mutation converts a stop codon to a tryptophan, predicted to result in the fusion of *hflD*, the lysogenization regulator, to *purB*. The *purB* gene encodes an adenylosuccinate lyase involved in purine biosynthesis. The Nal^r clone has a lower growth rate than the Str^r isolate when cultured in minimal media, a difference which is corrected upon addition of adenine or by genetic complementation with *purB*. Addition of adenine or genetic complementation also restored the invasion efficiency of the Nal^r clone. This report reconciles longstanding inconsistencies in phenotypic properties of an archetypal strain and provides both reassurance and cautions regarding intentional and unintentional evolution *in vitro*.

Escherichia coli was first established as a cause of diarrhea in the 1940s when serologically distinct strains were associated with outbreaks of severe and often fatal infections in infants (1, 2). A period of doubt regarding the virulence of these enteropathogenic *E. coli* (EPEC) strains came to a close in 1978 when it was reported that EPEC strains from the O142 and O127 serogroups, which were distinct from enterotoxigenic and enteroinvasive pathotypes, induced diarrhea in volunteers (3). The O127:H6 strain used in this study, E2348/69 (also known as E2348 or E2348-69), has since become the prototype EPEC strain used to study pathogenesis. As of April 2013, a search of Google Scholar (http://scholar.google.com/) using the term "E2348" yielded over 2,000 references (our unpublished observations).

Upon infection, EPEC colonizes the intestinal epithelium and dramatically alters the cells to which the bacteria adhere, causing effacement of the microvilli, rearrangement of the cytoskeleton, and the formation of actin-rich pedestals, a process known as attaching and effacing (A/E) (4–6). The ability to form A/E lesions is shared among EPEC strains, enterohemorrhagic *E. coli* (EHEC) strains, and strains of *Escherichia albertii* and *Citrobacter rodentium* (7–9). All attaching and effacing pathogens carry the locus of enterocyte effacement (LEE), a genetic element that harbors the genes encoding intimin, *eae*, the adhesin required for the A/E phenotype, and a type 3 secretion system (10). Typical EPEC (tEPEC) strains also carry large EPEC adherence factor (EAF) virulence plasmids that can be detected with gene probes or by PCR (11). The EAF plasmid contains the type IV bundle-forming pilus (*bfp*) gene cluster (12, 13), which is required for localized adherence of EPEC to epithelial cells (14) and for autoaggregation and contributes to virulence in volunteers (15). While EPEC is classified as an extracellular pathogen, it does have the ability to be internalized by epithelial cells *in vitro* (16, 17) and by enterocytes in the small intestine as seen in biopsy specimens from EPEC-infected infants (6, 18). The invasion efficiency is dependent on the virulence factors carried on the LEE and the EAF plasmid (19).

The complete genome sequence of strain E2348/69 was published in 2008 (20). This strain was first isolated in October 1969 from a 9-day-old boy amid an outbreak of infantile diarrhea afflicting 17 infants in a residential nursery in Taunton, United Kingdom, and was stored in lyophilized form at the Health Protection Agency (HPA), Colindale, London, United Kingdom (21, 22) (Claire Jenkins, HPA, personal communication). It has been reported that the sequenced strain was acquired from original stocks at the HPA and studied with minimal passages (20). This clone is resistant to streptomycin (Str^r) due to the presence of the *strAB* operon encoding a phosphotransferase on a small plasmid called pE2348-2 (23). In contrast, the E2348/69 strain used in most laboratories (14, 24, 25) is sensitive to streptomycin (Str^s) but resistant to nalidixic acid (Nal^r). The resistance to nalidixic acid was the result of intentional selection performed to facilitate recovery during volunteer studies at the Center for Vaccine Development (CVD), University of Maryland School of Medicine, at some time between the 1978 study (3) and 1985 (21). These differences suggest that there are at least three different clones referred to as E2348/69: (i) the original Str^r clone, (ii) an interme-

Received 19 June 2013 **Accepted** 26 July 2013

Published ahead of print 2 August 2013

Address correspondence to Michael S. Donnenberg, mdonnenb@medicine.umaryland.edu.

* Present address: Jean-Philippe Nougayrède, INRA USC 1360, Centre de Physiopathologie de Toulouse Purpan (CPTP), Toulouse, France.

doi:10.1128/JB.00704-13

TABLE 1 Strains and plasmids used in this study

Strain or plasmid	Description	Source or reference
Escherichia coli strains		
E2348/69 Str^r	Wild-type EPEC serotype O127:H6 resistant to streptomycin	Health Protection Agency, Colindale, London, United Kingdom
E2348/69 Nal^r	E2348/69 Str^r Δ*pE2348-2 gyrA*, *ftsK* Φ*hflD-purB* resistant to nalidixic acid	Center for Vaccine Development, University of Maryland
E2348/69 Str^s Nal^s	E2348/69 Str^r Δ*pE2348-2* sensitive to streptomycin and nalidixic acid	Center for Vaccine Development, University of Maryland
UMD753	E2348/69 Nal^r Δ*nleF*	31
UMD731	E2348/69 Nal^r Δ*escF*::*kan*	This study
E2348/69 Str-GFP	E2348/69 Str^r *att*Tn*7*::*gfpmut2*	This study
E2348/69 Nal-GFP	E2348/69 Nal^r *att*Tn*7*::*gfpmut2*	This study
Plasmids		
pXLW45	pGRG36 cloned with *gfpmut2* in front of the sigma70 promoter—allows integration into chromosomal *att*Tn*7* site	30
pBad24	Expression vector with pBAD promoter	Invitrogen
pSYN50	*hflD* cloned into pBad24	This study
pSYN56	*purB* cloned into pBad24	This study

diate clone either sensitive to both streptomycin and nalidixic acid or resistant to both antimicrobials, and (iii) the extant Nal^r clone. Adding further complexity to the cadre of E2348/69 clones, another group reported the presence of two small plasmids in their Str^r version of E2348/69, one identical to pE2348-2 and another that they designated p5217 (23).

During early studies of cellular invasion, it was noted by one of us that the Nal^r version of E2348/69 obtained from the CVD was less efficient in invading epithelial cells than other EPEC strains (unpublished data). A lyophilized vial of the Str^r strain was therefore obtained from the HPA, stored at −80°C after two passages on Luria-Bertani (LB) agar plates, and used in subsequent invasion studies to identify genes required for pathogenesis (17, 19). The genetic basis for this phenotypic difference has not previously been evaluated.

The pathogenesis of bacterial infections is frequently studied using mutants and complemented strains (26). One of the more popular procedures for creating mutants in *E. coli* is the one-step PCR method employing bacteriophage lambda recombinases (27). Despite its designation, as many as 15 passages are required between recovery of the parent from storage and storage of isolated colonies of the final mutant. Unintended mutations could occur during construction of these mutants, as has been reported using allelic exchange (28, 29).

The aim of this study was to determine using whole-genome sequencing and phenotypic assays the genetic differences and consequences thereof among different clones and mutants of the prototype EPEC strain.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The bacterial strains and plasmids used in this study are listed in Table 1, and the primers used are listed in Table 2. Overnight bacterial cultures were grown in LB broth with shaking (225 rpm) at 37°C from single colonies obtained on LB plates from frozen glycerol stocks. All growth assays were carried out in a 50-ml volume in 125-ml flasks with shaking (225 rpm) at 37°C. Growth assays conducted in M9 minimal medium and Dulbecco's modified Eagle's medium (DMEM)/F12 medium were initiated after overnight culture in LB by subculture with a 1:50 dilution, while growth assays were conducted using LB and a 1:100 dilution. Cultures were supplemented with ampicil-

TABLE 2 Primers used in this study

Name	Sequence (5′–3′)
Donne-481	TATTAACTTCTGAGGGAAATTTAATGAATTTATCTGTGTAGGCTGGAGCTGCTTCG
Donne-482	CCACACCAGTATCTTATTAGCAGAAATATCATTAACCATATGAATATCCTCCTTAG
HflD_F2	GGTACCCCTAGGGTGGCAAAGAATTACTATGACATCACC
HflD_R2	GTCGACGGCGCGCCTCACAACTCCGGGGTTAAATGA
PurB_F2	GGTACCCCTAGGATGGAATTATCCTCACTGACCGC
PurB_R2	GTCGACGGCGCGCCTTATTTCAGCTCATCAACCATGGTGAT
BfpE_F1	ATGAAAGAGAAATTAAACAGACTGCTATTC
BfpE_R1	AATCCCTGCCAGTGCGCC
MobA_F1	ATGATTGTTAAATTTCATGCCAGGGG
MobA_R1	TTACCATCCCAGCGAAGGG
SulII_F1	ATGAATAAATCGCTCATCATTTTCGGC
SulII_R1	TTAACGAATTCTTGCGGTTTCTTTCA
StrA_F1	GGTTCGATGAATTGCGGCT
StrA_R1	TCACGACCGAGCAA
StrB_F1	TGATCGACCTTGGGCGG
StrB_F1	GAGCTGTGGGGCGC
BfpC_F	ATGATAAAGAATAATCTTGGCGTAGC
BfpC_R	CTAAGCTCTCCCCCAAAAAACACCT
E2348C_0698_F1	ATGAAACGAACGCCTGTCC
E2348C_0698_R1	CTTCTGCCCCTGCGG
E2348C_1101_F1	CTAAGCGAGGTTCTGGTTTTCAAATT
E2348C_1101_R1	ATGAAGTATGTCTTTATTGAAAAACATCAGGC
E2348C_1103_F1	CAGAGTGAAATCTTTGCCCGA
E2348C_1103_R1	ACAGACTGCACGAAGCG
E2348C_0988_F1	TCAGAATGGACACGGCC
E2348C_0988_R1	ATGTCACTTGCAACCACAGTG
E2348C_0677_F1	GGATGAAAAAACTAACCTTTGAAATTCGAT
E2348C_0677_R1	TTTCCATTCGAGAGCTAGTCGC
E2348C_1947_F1	CTAAAGGTTTTGGCTTCATTACTCCG
E2348C_1947_R1	ACCTTTCTGACCGTCCTGAATTT
E2348C_2976_F1	ATGGCAAAAAACACATCTTGCG
E2348C_2976_R1	CGTATTCATGACGCCACCC
E2348C_4085_F1	GTGAATTTACTGACAGTGAGTACTGATCTC
E2348C_4085_R1	GATGTATGGCAGGATGGCG
E2348C_2375_F1	ATGAGCGACCTTGCGAGA
E2348C_2375_R1	ATCATCGATATACGCCAGACAACC
E2348C_0886_F1	GTACCACAGCCGCAGGATA
E2348C_0886_F1	CTCCATCTCGTTAACACACCAG
E2348C_1273_F1	GTGGCAAAGAATTACTATGACATCACC
E2348C_1273_R1	ACGGCGAGATCTTTAATGCCAT

lin at 200 μg/ml, adenine at 100 μg/ml, and L-arabinose at 0.2% (wt/vol) as required unless stated otherwise and monitored for absorbance at an optical density at 600 nm (OD_{600}) for 24 h. The growth rate constants were calculated by linear regression using Microsoft Excel from the logarithmic phase of the growth curves (from 2 to 6 h for cultures grown in DMEM/F13, 4 to 8 h for cultures grown in M9, and 1 to 4 h for cultures grown in LB). Reported constants represent the averages of the results of 3 to 6 trials.

Construction of strains and plasmids. Restriction enzymes (Fermentas), T4 DNA ligase (Roche), and Platinum Pfx DNA polymerase (Invitrogen) were used according to the manufacturers' instructions. The oligonucleotides used in this study were synthesized by Integrated DNA Technologies (IDT), and small-scale DNA sequencing was performed by the sequencing facility at the University of Maryland Biopolymer-Genomics Core. The Nal-green fluorescent protein (Nal-GFP) and Str-GFP strains were constructed using pXLW45 as described previously (30). UMD731, an *escF* mutant of the Nal^r strain, was constructed using the lambda recombinase procedure and primers Donne-481 and Donne-482 (Table 2) (27). Initial characterization of a prospective mutant revealed that it had lost the EAF plasmid. Therefore, chromosomal DNA from this mutant was purified, concentrated, and used as the DNA source for a second round of the procedure. Finally, the antibiotic cassette was removed, leaving a scar, as previously described (27). For plasmid constructions, the *hflD* and *purB* genes were amplified from the E2348/69 Str^r chromosomal DNA using primers HflD_F2 and HflD_R2 and primers PurB_F2 and PurB_R2, respectively (Table 2). These products were cloned into the pCR-blunt vector (Invitrogen), sequenced, and subcloned into pBad24 to generate plasmids pSYN56 and pSYN50.

Whole-genome sequencing. Genomic DNA from E2348/69 Nal^r and from a previously constructed *nleF* mutant strain, UMD753 (31), were isolated from overnight cultures grown in LB broth (Difco) using a Sigma GenElute kit (Sigma-Aldrich). The genomes were sequenced at the Institute for Genome Sciences Genome Resource Center (http://www.igs.umaryland.edu/resources/grc/index.php) on the Illumina HiSeq2500 sequencer using paired-end libraries with 300-bp insertions. The genomes were assembled using the Velvet assembly program (32) with kmer values determined using VelvetOptimiser v2.1.4 (http://bioinformatics.net.au/software.velvetoptimiser.shtml). Contigs ≥ 500 bp were used for further analysis.

Single nucleotide polymorphisms (SNPs) were identified in the draft genome sequences of the Nal^r clone and UMD753 compared to that of the Str^r strain previously described (33–35). Briefly, SNPs were identified in the draft assemblies of the Nal^r strain and UMD753 compared to the closed chromosome (NC_011601.1) and the plasmids (NC_011603.1 and NC_011602.1) of the Str^r strain using MUMmer (36). The SNPs were verified by PCR amplifying and sequencing the SNP and the surrounding region to at least 140 bp in the Nal^r clone, in UMD753, and in the Str^r clone from our collection using the primers listed in Table 2.

Flow cytometry analysis. GFP-labeled Str^r (Str-GFP) and Nal^r (Nal-GFP) clones were analyzed using flow cytometry to detect cell division and/or chromosome segregation defects (37). To analyze cell division, overnight cultures were subcultured in 3 ml LB with a 1:100 dilution in 15 ml tubes with and without 15 μg/ml ampicillin and grown for 3 h with shaking at 37°C. Cells were washed twice for 1 min at 14,000 × *g* in phosphate-buffered saline (PBS) and resuspended in PBS for analysis. To analyze chromosome segregation, cells were grown as described above, and 500 μl of cultured cells was washed twice for 1 min at 14,000 × *g* in 50 mM HEPES buffer and incubated with 20 μM SYTO 17 (Molecular Probes) for 30 min at room temperature, washed twice for 1 min at 14,000 × *g* in 50 mM HEPES, and resuspended in 50 mM HEPES for analysis. Flow cytometry analyses were performed on a BD LSRII flow cytometer (Becton, Dickinson). The data from at least 10^4 cells were collected using FACSDiva Software (Becton, Dickinson).

Invasion assay. Invasion of Hep-2 cells by the various EPEC clones was determined as described previously (17). When required, 100 μg/ml

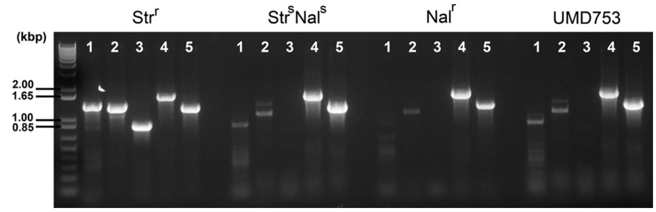

FIG 1 Agarose (1%) gel electrophoresis of PCR products amplified from E2348/69 plasmids. Lanes 1 to 5: *strA* (1.32 kbp), *strB* (1.26 kbp), *sulII* (0.816 kbp), *mobA* (1.55 kbp), and *bfpC* (1.2 kbp). The *strA*, *strB*, and *sulII* genes are on the pE2348-2 plasmid. The *mobA* and *bfpC* genes are on plasmid p5217 and the EAF plasmid, respectively. The various E2348/69 clones are indicated above the gel. Both plasmid p5217 and the EAF plasmid were present in all strains, whereas pE2348/69 was present only in the Str^r clone.

of adenine was added to the assay during the 3-h infection and 0.2% L-arabinose was added both to the logarithmic-phase bacterial cultures and during the 3-h infection period.

Nucleotide sequence accession numbers. The draft genome assemblies of E2348/69 Nal^r and E2348/69 Nal^r Δ*nleF* (UMD753) have been deposited in GenBank under accession numbers ASZR00000000 and ASZS00000000, respectively.

RESULTS

Sequencing results and comparative analysis of genomes. The draft genome sequence of the E2348/69 Nal^r strain is approximately 4.9 Mb and is assembled into 130 contigs that are ≥500 bp. Comparison of the draft genome sequences of E2348/69 Nal^r and UMD753 revealed no differences in the assembled genomes except for the *nleF* deletion mutation created using one-step PCR mutagenesis (31). Comparison of the E2348/69 Nal^r sequence to the published sequence of E2348/69 Str^r revealed the presence of the p5217 plasmid (23) and the loss of the pE2348-2 plasmid in both E2348/69 Nal^r and UMD753 (Fig. 1). Comparative analysis also identified 19 genetic differences located on the EAF plasmid. Eighteen of these changes were present either on repeat elements or in genes encoding transposases. As these differences likely represent slightly different alleles of these multicopy genes, they were not further investigated. The last difference was within the *bfpE* gene. PCR and sequencing of this gene confirmed a single synonymous difference (Table 3).

There were 23 genetic differences identified in the chromosomal DNA of E2348/69 Nal^r (Table 3) compared to that of the published E2348/69 Str^r strain (20). PCR and sequencing of these regions revealed that eight were sequencing errors, eight were synonymous mutations in multicopy genes, four were sequence differences between the published genome and the Str^r and Nal^r strains in our collection, and three were nonsynonymous mutations present only in the Nal^r clone (Table 3).

Phenotypic changes due to changes in genotype. The sequencing of the Nal^r clone and *nleF* mutant revealed a nonsynonymous point mutation in comparison to the Str^r reference genome, resulting in a substitution of serine for leucine at position 83 (S83L) in the DNA gyrase A subunit. This mutation results in quinolone resistance in *E. coli* (38). Thus, we conclude that S83L is the cause of the nalidixic acid resistance in E2348/69 Nal^r.

FtsK is a DNA translocase that is essential for the coordination of cell division and chromosome segregation (39). The nonsynonymous mutation of the *ftsK* gene (G969S) results in a conservative substitution in a loop between two beta sheets of the C-termi-

TABLE 3 Validation of SNPs

Product(s)	Position	Str^r (NCBI)	Str^r (this study)	Nal^r	UMD753
bfpE (E2348_P1_009)	72	AGG (R)	AGG (R)	AGA (R)	AGA (R)
Multicopy genes					
Bacteriophage 21 head gene operon	381	GGG (G)	GGG (G) GGA (G)	GGG (G) GGA (G)	GGG (G) GGA (G)
(E2348C_0698)	393	TTT (F)	TTT (F) TTC (F)	TTT (F)	TTT (F) TTC (F)
Transposase Orf2 of ISEc17 (E2348C_1011)	543	CTA (L)	CTA (L) CTG (L)	CTA (L) CTG (L)	CTA (L) CTG (L)
	272	CAT (H)	CAT (H) CGT (R)	CAT (H) CGT (R)	CAT (H) CGT (R)
Transposase Orf3 of IS66 family insertion	832	AGT (S)	AGT (S) GGT (G)	AGT (S) GGT (G)	AGT (S) GGT (G)
sequence element (E2348C_1103)	861	GCT (A)	GCT (A)	GCT (A)	GCT (A)
Predicted protein (bacteriophage) (E2348C_0988)	84	GGC (G)	GGT (G)	GGT (G) GGC (G)	GCC (A) GGT (G)
Bacteriophage lambda NinB recombination proteins	33	CCA (P)	CCA (P) CCG(P)	CCG(P)	CCA (P) CCG(P)
Sequencing differences					
Stress protein, member of the CspA family	120	GGT (G)	GGC (G)	GGC (G)	GGC (G)
(E2348C_1947)	102	TCT (S)	TCC (S)	TCC (S)	TCC (S)
Hydrogenase maturation protein (HypF) (E2348C_2976)	94	TAA (STOP)	CAA (Q)	CAA (Q)	CAA (Q)
UDP-GlcNAc: undecaprenylphosphate GlcNAc-1-phosphate transferase (E2348C_4085)	440	GAG (E)	GCG (A)	GCG (A)	GCG (A)
Nonsynonymous mutations					
DNA gyrase (type II topoisomerase), subunit A (E2348C_2375)	248	TCG (S)	TCG (S)	TTG (L)	TTG (L)
DNA-binding membrane protein FtsK required for chromosome resolution and partitioning (E2348C_0886)	2905	GGC (G)	GGC (G)	AGC (S)	AGC (S)
Predicted lysogenization regulator (*hflD*) fused to *purB* (E2348C_1273)	642	TGA (STOP)	TGA (STOP)	TGG (W)	TGG (W)

nal motor domain, as determined from the crystal structure deposited in NCBI (Molecular Modeling Database [MMDB] identification no. 41228). To determine if the mutation in *ftsK* affected cell division and/or chromosomal segregation in the Nal^r clone, both the Str^r and Nal^r clones were cultured in LB and monitored for alterations in growth rate and cellular morphology using GFP. There was no significant difference between the growth constants of the Str-GFP and Nal-GFP clones when cultured in LB (Table 4). Furthermore, the presence of *gfp* at the *att*Tn*7* site did not alter the growth constant compared to that of the Str^r strain (Table 4). Microscopic examination and flow cytometry analysis of Str-GFP and Nal-GFP cells did not show the cell elongation that would be expected when a region important to FtsK function is perturbed (37) (Fig. 2A). Nucleic acid staining with SYTO 17 did not reveal chromosomal segregation differences between the Str-

TABLE 4 Growth rate constants

	Growth rate (OD^{600} h^{-1}) in indicated medium		
Clone of E2349/69	LB	M9	DMEM/F12
Str-GFP	0.73 ± 0.01		
Nal-GFP	0.77 ± 0.04		
Str^s Nal^s		1.12 ± 0.03	
Str^r	0.70 ± 0.01	1.31 ± 0.14	2.03 ± 0.03
Nal^r		0.74 ± 0.03[a]	1.23 ± 0.08
Nal^r + adenine		1.18 ± 0.09	
Nal^r (pBad24)		1.41 ± 0.04	
Nal^r (pSYN56)		1.35 ± 0.03	
Nal^r (pSYN50)		1.60 ± 0.09[b]	
Str^r + L-arabinose		1.83 ± 0.08	
Nal^r + L-arabinose		1.62 ± 0.02	

[a] $P \leq 0.02$ (compared to Str^r and Nal^r + adenine).
[b] $P \leq 0.02$ (compared to Nal^r [pBad24] and Nal^r [pSYN56]).

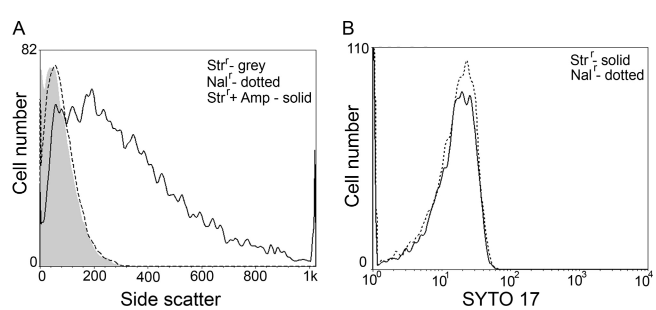

FIG 2 The G969S mutation in the *ftsK* gene does not affect cell division or chromosomal segregation. (A) Representative histogram of the side scatter of logarithmic-phase GFP-labeled Str^r, Nal^r, and Str^r plus 15 μg/ml ampicillin cells. In contrast to the results seen with elongated cells grown in the presence of 15 μg/ml of ampicillin as a positive control, side scatter from Nal^r cells was indistinguishable from that of Str^r cells. (B) Representative histogram of nucleic acid staining of logarithmic-phase GFP-positive Str^r and Nal^r cells. There is no evidence of increased DNA staining of the Nal^r cells to suggest a segregation defect due to the mutation in the *ftsK* gene. Experiments were performed in duplicate and repeated at least 3 times with similar results.

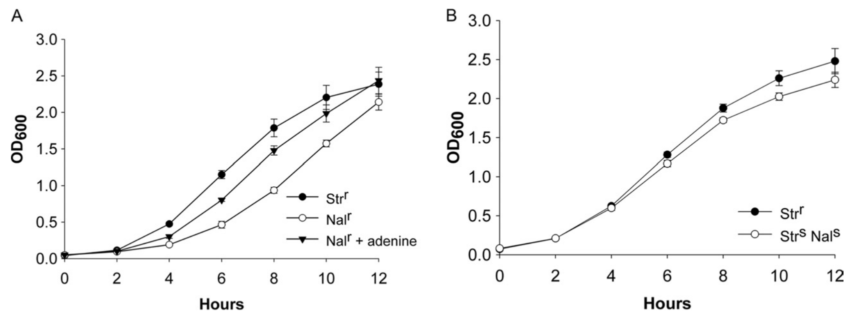

FIG 3 Growth curve of E2348/69 clones in M9 media. (A) The Str^r clone had a significantly higher growth rate constant than the Nal^r clone which was restored upon addition of adenine or by genetic complementation with *purB* (data not shown); however, the number of bacteria in the Nal^r plus adenine cultures at each time point remained lower than the number in the Str^r cultures. (B) The presence of the pE2348-2 plasmid conferred a growth advantage during the logarithmic phase, as seen from the Str^r cultures with the pE2348-2 plasmid compared to the Str^s Nal^s cultures without the pE2348-2 plasmid. Experiments were performed in duplicate and repeated at least 3 times; error bars indicate standard errors of the means. Student's paired *t* test with a two-tailed distribution showed a significant difference between the Str^r and Nal^r plus adenine cultures between h 2 and h 8 but showed a significant difference between the Str^r and Str^s Nal^s cultures only between h 8 and h 10.

GFP and Nal-GFP cells (Fig. 2B). We conclude that the *ftsK* mutation in the Nal^r strain does not result in FtsK dysfunction.

The nonsynonymous mutation in the *hflD* gene converts the stop codon to a tryptophan, resulting in a fusion with the downstream *purB* gene. HflD is a negative regulator of bacteriophage lambda lysogenization (40), while the *purB* product is an adenylosuccinate lyase required for purine biosynthesis (41). To determine if the fusion of the *hflD* gene to *purB* affected the synthesis of adenine, both the Str^r and Nal^r clones were cultured in M9 minimal media. The Str^r clone grew significantly faster than the Nal^r clone (Table 4). This growth defect was restored to a rate similar to that of the Str^r clone upon addition of adenine or by genetic complementation with the *purB* gene (Nal^r-pSYN50). In contrast, addition of the *hflD* gene (Nal^r-pSYN56) had no significant effect on growth (Table 4). Growth rate constants of strains complemented with plasmids were higher than those of strains without complementation, a finding we attribute to the addition of 0.2% L-arabinose to the minimal media of the former, which may serve as an additional carbon source (Table 4). While the growth rate constant during the logarithmic phase was restored upon complementation, the number of bacteria in the Nal^r cultures at each time point was still lower than that in the Str^r cultures (Fig. 3A). We hypothesize that this difference may be due to the presence of pE2348-2 in the Str^r clone. Previous work by Enne et al. has shown that the presence of a similar plasmid, p9123, which contains the *strA*, *strB*, and *sulII* genes confers a growth advantage (42). Growth of our intermediate Str^s Nal^s clone showed that while the presence of the pE2348-2 plasmid confers a certain growth advantage, it does not entirely reflect the growth pattern of the Nal^r plus adenine strain (Fig. 3B).

Invasion efficiency of various clones. Most experiments using an EPEC infection model are carried out for 3 h, which falls in the logarithmic growth phase, and in DMEM/F12 medium which is not supplemented with nucleotides (17, 19, 24, 25). Similar to the growth in minimal media, the growth rate of the Nal^r clone was compromised in the logarithmic phase in DMEM/F12. We sought to establish whether the *hflD-purB* gene fusion was responsible for the defect in epithelial cell invasion. Figure 4 shows that the ability of the Nal^r clone to invade epithelial cells was compromised compared to that of the Str^r clone. Invasion efficiency was restored to the Str^r strain upon addition of adenine or by genetic complementation with the *purB* but not the *hflD* gene.

DISCUSSION

We owe most of our current understanding of the mechanisms by which EPEC causes disease to studies using a single strain, E2348/69. However, like many archetypal strains, E2348/69 has a long history of passage and storage under various conditions in different laboratories, with attendant uncertainty regarding genomic fidelity. Furthermore, various methods of mutagenesis and phenotypic differences among variants of this strain present the potential for confusion regarding the results of different studies. The current report resolves some of this uncertainty by documenting the genetic changes that have occurred in multiple versions of the strain during more than 40 years of passage and storage and has general implications for studies of bacterial pathogenesis using genetic techniques.

One of the more reassuring results from this study is the find-

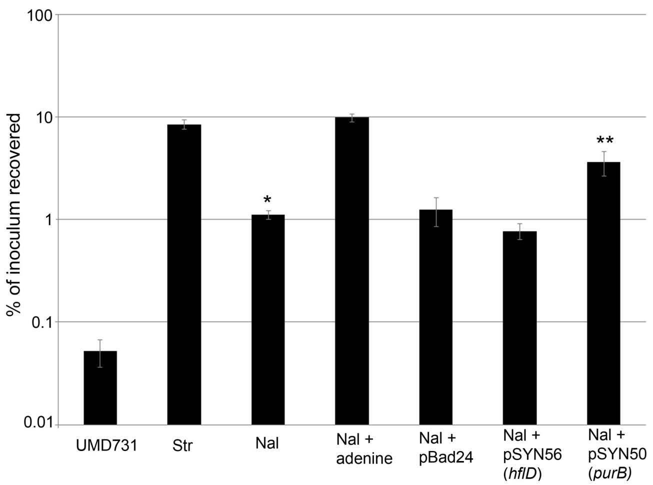

FIG 4 The HEp-2 cell invasion efficiency of the Nal^r clone was compromised due to the fusion of the *purB* gene to the *hflD* gene. The Nal^r clone was less efficient at HEp-2 cell invasion than the Str^r clone. This deficiency was restored upon addition of adenine or by genetic complementation with the *purB* gene (pSYN56) but was unaffected by complementation with the *hflD* gene (pSYN50). UMD731, an *escF* mutant, was used as a negative control. Experiments were performed in duplicate and repeated at least 3 times; error bars indicate standard errors of the means. *, $P \leq 0.001$ (compared to Str^r and Nal^r plus adenine); **, $P \leq 0.05$ (compared to Nal^r plus pBad24 and Nal^r plus pSYN56).

ing that no unintended mutations occurred during the construction of an *nleF* mutant strain, UMD753. Our experience constructing another mutant strain, UMD731, which has a deletion in the *escF* gene encoding the T3SS needle (24) and served as a control in the invasion studies described here, had caused us to question the fidelity of the procedure. Our analysis of strain UMD753 suggests that this procedure does not readily result in unintended mutations.

Additionally, this work confirms the presence in the Str^r clone of the previously described 5.2-kb p5217 plasmid containing the *mob* genes (23), which was not included in the published sequence (20). We used primers specific to the *mobA* gene and verified that p5217 was present in our stocks of both the Str^r and the Nal^r clones (Fig. 1). It is not clear whether the plasmid in the published reference Str^r genome was lost during the minimal passage prior to sequencing or whether its presence was missed due to similarities with pE2348-2. Using primers specific to the *strA*, *strB*, and *sulII* genes, we verified that the pE2348-2 plasmid was absent in our Nal^r clones but was present in our Str^r clone (Fig. 1). The *strAB* operon contributes to streptomycin resistance in many bacterial species, including E2348/69 (23); thus, we attribute the streptomycin sensitivity of the Nal^r clones to the loss of the pE2348-2 plasmid.

Apart from the small cryptic plasmids, the draft genome sequence of the Nal^r strain identified 19 potential SNPs on the EAF plasmid and 23 on the chromosomal DNA in comparison to the published Str^r genome sequence (Table 3). Only three of these SNPS were nonsynonymous changes between both the published Str^r genome and our Str^r clone compared to the Nal^r clones in single-copy genes: one each in the *gyrA*, *ftsK*, and *hflD* genes (Table 3).

Substitutions at position 83 in the DNA gyrase A subunit give rise to resistance to nalidixic acid and other quinolones (38). Thus, we attribute the cause of nalidixic acid resistance in the Nal^r clones to this mutation. Subsequently, we checked for nalidixic acid and streptomycin resistance in all the E2348/69 strains stocked at various times over the years in the CVD collection. We were able to identify an E2348/69 clone frozen in 1977 that was sensitive to both streptomycin and nalidixic acid. PCR amplification of *gyrA* in this clone showed it to be identical to the DNA gyrase A subunit of the Str^r clones, i.e., without the S83L mutation seen in the Nal^r clones. These results indicate a stepwise progression in the evolution of the Nal^r clone that begins with the original Str^r isolate, includes an intermediate Str^s and Nal^s clone after loss of pE2348-2, and concludes with the extant Nal^r clone. An alternative Str^r and Nal^r intermediate has not be recovered and is unlikely to have been made, as resistance to streptomycin would have precluded the rationale for selecting nalidixic acid resistance.

We also noted that the Nal^r clones have a conservative substitution in the *ftsK* gene, encoding an essential cell division and chromosome segregation protein. Neither the Str^r nor Nal^r clones showed any growth defects in rich medium (Table 4). FtsK recruits other divisome proteins, and perturbation of this process can lead to a filamentous phenotype (37). Flow cytometry analysis of Str-GFP and Nal-GFP cells did not show a filamentous phenotype, whereas cells grown in the presence of 15 μg/ml ampicillin did (Fig. 2A). Both the Str-GFP and Nal-GFP cells were also stained with SYTO 17 to determine whether there were any chromosomal segregation defects, but no difference in DNA staining

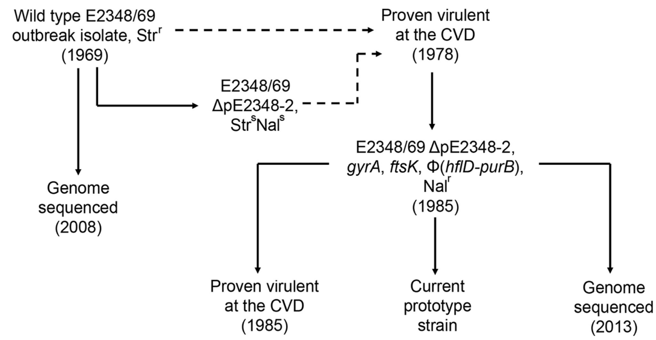

FIG 5 Flow chart depicting the *in vitro* evolution of E2348/69. The exact clone used for the virulence study in 1978 is unknown (shown with dotted lines). The extant Nal^r E2348/69 clone used as the "wild type" has lost the pE2348-2 plasmid and has 3 nonsynonymous mutations.

was found. These data suggest that the mutation in the *ftsK* gene is unlikely to have a significant effect on the function of the protein.

The other nonsynonymous mutation that occurred during the evolution of the Nal^r clone converted the stop codon of the *hflD* gene to a tryptophan, causing it to fuse to the downstream *purB* gene. Interestingly, mutations in the *purB* gene have been shown to cause low-level resistance to nalidixic acid (43). This mutation may have arisen while selecting for Nal^r clone for volunteer studies. However, due to the dominant *gyrA* mutation, we did not attempt to verify whether the *hflD-purB* fusion contributed to nalidixic acid resistance in the Nal^r clone. PurB is an adenylosuccinate lyase required in the IMP synthesis pathway to purine biosynthesis (44). To investigate if the *hflD-purB* fusion affected adenylosuccinate lyase activity, we cultured the Str^r and Nal^r clones in M9 minimal and DMEM/F12 media and found that the Nal^r clone has a significantly lower growth rate constant than the Str^r clone (Table 4). The lower growth rate was restored to the same level as the Str^r clone growth rate upon addition of external adenine or via genetic complementation with the *purB* gene on a plasmid (Table 4). These results confirm that the *hflD-purB* fusion compromised the cells' ability to synthesize purines. Since the *hflD-purB* fusion has a significant effect on the growth of the Nal^r clone, we examined whether this mutation was also responsible for the previously observed deficiency in epithelial cell invasion, as these experiments are carried out in media that does not contain purine supplements. Figure 4 shows that the invasion efficiency of the Nal^r clone is lower than that of the Str^r clone. Furthermore, invasion efficiency is restored upon addition of adenine or by overexpressing *purB* but not by overexpressing *hflD*.

The regions around the three nonsynonymous mutations found in the Nal^r clone were also analyzed in the Str^s Nal^s clone and were found to be identical to those found in the Str^r clone. These results led to the reconstructed history displayed in Fig. 5.

In conclusion, we sequenced two variants of the E2348/69 strain and found that the archetypal E2348/69 strain has undergone a number of genomic changes over time, only one of which was intentional. While these changes do not impair the virulence of E2348/69 in volunteer studies (21), the fusion of *hflD* to *purB* compromises the ability of the strain to grow in minimal media, thus making it less invasive under these conditions. Our results highlight the pitfalls of working with archetypal strains whose pedigrees are incompletely characterized.

ACKNOWLEDGMENTS

We are grateful to Xiaolin Wang for making the Nal-GFP strain. We thank Claire Jenkins from the Health Protection Agency for providing patient data for the original E2348/69 isolate and Jane Michalski for providing the different clones of E2348/69 stocked at the CVD.

Research reported in this publication was supported by National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health, Enteric Research Investigational Network (ERIN), under award U19AI090873.

REFERENCES

1. **Bray J.** 1945. Isolation of antigenically homogeneous strains of *Bact. coli neapolitanum* from summer diarrhoea of infants. J. Pathol. Bacteriol. **57:** 239–247.
2. **Giles C, Sangster G, Smith J.** 1949. Epidemic gastroenteritis of infants in Aberdeen during 1947. Arch. Dis. Child. **24:**45–53.
3. **Levine MM, Bergquist EJ, Nalin DR, Waterman DH, Hornick RB, Young CR, Sotman S, Rowe B.** 1978. *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. Lancet **i:**1119–1122.
4. **Knutton S, Lloyd DR, McNeish AS.** 1987. Adhesion of enteropathogenic *Escherichia coli* to human intestinal enterocytes and cultured human intestinal mucosa. Infect. Immun. **55:**69–77.
5. **Moon HW, Whipp SC, Argenzio RA, Levine MM, Giannella RA.** 1983. Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. Infect. Immun. **41:**1340–1351.
6. **Ulshen MH, Rollo JL.** 1980. Pathogenesis of *Escherichia coli* gastroenteritis in man—another mechanism. N. Engl. J. Med. **302:**99–101.
7. **Huys G, Cnockaert M, Janda JM, Swings J.** 2003. *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. Int. J. Syst. Evol. Microbiol. **53:**807–810.
8. **Frankel G, Phillips AD, Novakova M, Field H, Candy DC, Schauer DB, Douce G, Dougan G.** 1996. Intimin from enteropathogenic *Escherichia coli* restores murine virulence to a *Citrobacter rodentium eaeA* mutant: induction of an immunoglobulin A response to intimin and EspB. Infect. Immun. **64:**5315–5325.
9. **Tzipori S, Wachsmuth IK, Chapman C, Birner R, Brittingham J, Jackson C, Hogg J.** 1986. The pathogenesis of hemorrhagic colitis caused by *Escherichia coli* O157:H7 in gnotobiotic pigs. J. Infect. Dis. **154:**712–716.
10. **Elliott SJ, Wainwright LA, McDaniel TK, Jarvis KG, Deng Y, Lai L-C, McNamara BP, Donnenberg MS, Kaper JB.** 1998. The complete sequence of the locus of enterocyte effacement (LEE) of enteropathogenic *E. coli* E2348/69. Mol. Microbiol. **28:**1–4.
11. **Kaper JB.** 1996. Defining EPEC. Rev. Microbiol. Sao Paulo **27**(Suppl 1):130–133.
12. **Stone KD, Zhang H-Z, Carlson LK, Donnenberg MS.** 1996. A cluster of fourteen genes from enteropathogenic *Escherichia coli* is sufficient for biogenesis of a type IV pilus. Mol. Microbiol. **20:**325–337.
13. **Sohel I, Puente JL, Ramer SW, Bieber D, Wu C-Y, Schoolnik GK.** 1996. Enteropathogenic *Escherichia coli*: identification of a gene cluster coding for bundle-forming pilus morphogenesis. J. Bacteriol. **178:**2613–2628.
14. **Donnenberg MS, Girón JA, Nataro JP, Kaper JB.** 1992. A plasmid-encoded type IV fimbrial gene of enteropathogenic *Escherichia coli* associated with localized adherence. Mol. Microbiol. **6:**3427–3437.
15. **Bieber D, Ramer SW, Wu CY, Murray WJ, Tobe T, Fernandez R, Schoolnik GK.** 1998. Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. Science **280:**2114–2118.
16. **Andrade JR, Da Veiga VF, De Santa Rosa MR, Suassuna I.** 1989. An endocytic process in HEp-2 cells induced by enteropathogenic *Escherichia coli*. J. Med. Microbiol. **28:**49–57.
17. **Donnenberg MS, Donohue-Rolfe A, Keusch GT.** 1989. Epithelial cell invasion: an overlooked property of enteropathogenic *Escherichia coli* (EPEC) associated with the EPEC adherence factor. J. Infect. Dis. **160:**452–459.
18. **Fagundes-Neto U, Freymuller E, Gatti MSV, Schmitz LG, Scaletsky I.** 1995. Enteropathogenic *Escherichia coli* O111ab:H2 penetrates the small bowel epithelium in an infant with acute diarrhoea. Acta Paediatr. **84:**453–455.
19. **Donnenberg MS, Calderwood SB, Donohue-Rolfe A, Keusch GT, Kaper JB.** 1990. Construction and analysis of Tn*phoA* mutants of enteropathogenic *Escherichia coli* unable to invade HEp-2 cells. Infect. Immun. **58:** 1565–1571.
20. **Iguchi A, Thomson NR, Ogura Y, Saunders D, Ooka T, Henderson IR, Harris D, Asadulghani M, Kurokawa K, Dean P, Kenny B, Quail MA, Thurston S, Dougan G, Hayashi T, Parkhill J, Frankel G.** 2009. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. J. Bacteriol. **191:**347–354.
21. **Levine MM, Nataro JP, Karch H, Baldini MM, Kaper JB, Black RE, Clements ML, O'Brien AD.** 1985. The diarrheal response of humans to some classic serotypes of enteropathogenic *Escherichia coli* is dependent on a plasmid encoding an enteroadhesiveness factor. J. Infect. Dis. **152:** 550–559.
22. **Taylor J.** 1970. Infectious infantile enteritis, yesterday and today. Proc. R. Soc. Med. **63:**1297–1301.
23. **Handford CL, Stang CT, Raivio TL, Dennis JJ.** 2009. The contribution of small cryptic plasmids to the antibiotic resistance of enteropathogenic *Escherichia coli* E2348/69. Can. J. Microbiol. **55:**1229–1239.
24. **Wilson RK, Shaw RK, Daniell S, Knutton S, Frankel G.** 2001. Role of EscF, a putative needle complex protein, in the type III protein translocation system of enteropathogenic *Escherichia coli*. Cell. Microbiol. **3:**753–762.
25. **Kenny B, DeVinney R, Stein M, Reinscheid DJ, Frey EA, Finlay BB.** 1997. Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. Cell **91:**511–520.
26. **Falkow S.** 1988. Molecular Koch's postulates applied to microbial pathogenicity. Rev. Infect. Dis. **10**(Suppl 2):S274–S276.
27. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. **97:**6640–6645.
28. **Johnson JR, Lockman HA, Owens K, Jelacic S, Tarr PI.** 2003. High-frequency secondary mutations after suicide-driven allelic exchange mutagenesis in extraintestinal pathogenic *Escherichia coli*. J. Bacteriol. **185:** 5301–5305.
29. **Ferrières L, Hemery G, Nham T, Guerout AM, Mazel D, Beloin C, Ghigo JM.** 2010. Silent mischief: bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. J. Bacteriol. **192:**6418–6427.
30. **Sharma M, Ingram DT, Patel JR, Millner PD, Wang X, Hull AE, Donnenberg MS.** 2009. A novel approach to investigate the uptake and internalization of *Escherichia coli* O157:H7 in spinach cultivated in soil and hydroponic medium. J. Food Prot. **72:**1513–1520.
31. **Blasche S, Mörtl M, Steuber H, Siszler G, Nisa S, Schwarz F, Lavrik I, Gronewold TMA, Maskos K, Donnenberg MS, Ullmann D, Uetz P, Kögl M.** 2013. The *E. coli* effector protein NleF is a caspase inhibitor. PLoS One **8:**e58937.
32. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18:**821–829.
33. **Rasko DA, Webster DR, Sahl JW, Bashir A, Boisen N, Scheutz F, Paxinos EE, Sebra R, Chin CS, Iliopoulos D, Klammer A, Peluso P, Lee L, Kislyuk AO, Bullard J, Kasarskis A, Wang S, Eid J, Rank D, Redman JC, Steyert SR, Frimodt-Moller J, Struve C, Petersen AM, Krogfelt KA, Nataro JP, Schadt EE, Waldor MK.** 2011. Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. N. Engl. J. Med. **365:**709–717.
34. **Eppinger M, Mammel MK, LeClerc JE, Ravel J, Cebula TA.** 2011. Genomic anatomy of *Escherichia coli* O157:H7 outbreaks. Proc. Natl. Acad. Sci. U. S. A. **108:**20142–20147.
35. **Rasko DA, Worsham PL, Abshire TG, Stanley ST, Bannan JD, Wilson MR, Langham RJ, Decker RS, Jiang L, Read TD, Phillippy AM, Salzberg SL, Pop M, Van Ert MN, Kenefic LJ, Keim PS, Fraser-Liggett CM, Ravel J.** 2011. *Bacillus anthracis* comparative genome analysis in support of the Amerithrax investigation. Proc. Natl. Acad. Sci. U. S. A. **108:**5027–5032.
36. **Delcher AL, Phillippy A, Carlton J, Salzberg SL.** 2002. Fast algorithms for large-scale genome alignment and comparison. Nucleic Acids Res. **30:**2478–2483.
37. **Grenga L, Luzi G, Paolozzi L, Ghelardini P.** 2008. The *Escherichia coli* FtsK functional domains involved in its interaction with its divisome protein partners. FEMS Microbiol. Lett. **287:**163–167.
38. **Ruiz J.** 2003. Mechanisms of resistance to quinolones: target alterations,

decreased accumulation and DNA gyrase protection. J. Antimicrob. Chemother. **51:**1109–1117.
39. **Sherratt DJ, Arciszewska LK, Crozat E, Graham JE, Grainge I.** 2010. The *Escherichia coli* DNA translocase FtsK. Biochem. Soc. Trans. **38:**395–398.
40. **Kihara A, Akiyama Y, Ito K.** 2001. Revisiting the lysogenization control of bacteriophage λ. Identification and characterization of a new host component, HflD. J. Biol. Chem. **276:**13695–13700.
41. **Jung SC, Smith CL, Lee KS, Hong ME, Kweon DH, Stephanopoulos G, Jin YS.** 2010. Restoration of growth phenotypes of *Escherichia coli* DH5α in minimal media through reversal of a point mutation in *purB*. Appl. Environ. Microbiol. **76:**6307–6309.
42. **Enne VI, Bennett PM, Livermore DM, Hall LM.** 2004. Enhancement of host fitness by the sul2-coding plasmid p9123 in the absence of selective pressure. J. Antimicrob. Chemother. **53:**958–963.
43. **Helling RB, Adams BS.** 1970. Nalidixic acid-resistant auxotrophs of *Escherichia coli*. J. Bacteriol. **104:**1027–1029.
44. **Zhang Y, Morar M, Ealick SE.** 2008. Structural biology of the purine biosynthetic pathway. Cell. Mol. Life Sci. **65:**3699–3724.